

Subscribe (Full Service) Register (Limited Service, Free) Login

Search: The ACM Digital Library The Guide

+"(word combination" or sentence or phrase or paragraph or v

HE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used: word combination or sentence or phrase or paragraph or words compute corpus segments Found 42 of

210,707

Sort results

relevance

Save results to a Binder 2 Search Tips

Try an Advanced Search

Try this search in The ACM Guide

by Display results

expanded form *

Open results in a new

window

Results 1 - 20 of 42

Result page: 1 2 3

Relevance scale

Broad coverage paragraph segmentation across languages and domains

Caroline Sporleder, Mirella Lapata

July 2006 ACM Transactions on Speech and Language Processing (TSLP), Volume 3 Issue

Publisher: ACM Press

Full text available: pdf(300.91 KB) Additional Information: full citation, abstract, references, index terms

This article considers the problem of automatic paragraph segmentation. The task is relevant for speech-to-text applications whose output transcipts do not usually contain punctuation or paragraph indentation and are naturally difficult to read and process. Textto-text generation applications (e.g., summarization) could also benefit from an automatic paragaraph segementation mechanism which indicates topic shifts and provides visual targets to the reader. We present a paragraph segmentation mod ...

Keywords: Machine learning, paragraph breaks, segmentation, summarization

A compression-based algorithm for Chinese word segmentation

W. J. Teahan, Rodger McNab, Yingying Wen, Ian H. Witten September 2000 Computational Linguistics, Volume 26 Issue 3

Publisher: MIT Press

Full text available:

Publisher Site

pdf(1.34 MB) Additional Information: full citation, abstract, references, citings

Chinese is written without using spaces or other word delimiters. Although a text may be thought of as a corresponding sequence of words, there is considerable ambiguity in the placement of boundaries. Interpreting a text as a sequence of words is beneficial for some information retrieval and storage tasks: for example, fulltext search, word-based compression, and keyphrase extraction. We describe a scheme that infers appropriate positions for word boundaries using an adaptive language model that ...

A computational grammar of discourse-neutral prosodic phrasing in English

J. Bachenko, E. Fitzpatrick

September 1990 Computational Linguistics, Volume 16 Issue 3

Publisher: MIT Press

Full text available: Additional Information: full citation, abstract, references, citings

Publisher Site

We describe an experimental text-to-speech system that uses information about syntactic constituency, adjacency to a verb, and constituent length to determine prosodic phrasing for synthetic speech. A central goal of our work has been to characterize "discourse neutral" phrasing, i.e. sentence-level phrasing patterns that are independent of discourse semantics. Our account builds on Bachenko et al. (1986), but differs in its treatment of clausal structure and predicate-argument relations. Result ...

The disambiguation of nominalizations

Maria Lapata

September 2002 Computational Linguistics, Volume 28 Issue 3

Publisher: MIT Press

Full text available: pdf(471.69 KB)

Additional Information: full citation, abstract, references, citings, index terms

This article addresses the interpretation of nominalizations, a particular class of compound nouns whose head noun is derived from a verb and whose modifier is interpreted as an argument of this verb. Any attempt to automatically interpret nominalizations needs to take into account: (a) the selectional constraints imposed by the nominalized compound head, (b) the fact that the relation of the modifier and the head noun can be ambiguous, and (c) the fact that these constraints can be easily overr ...

Annotations and tools for an activity based Spoken Language Corpus

Jens Allwood, Leif Gröngvist

September 2001 Proceedings of the Second SIGdial Workshop on Discourse and Dialogue - Volume 16

Publisher: Association for Computational Linguistics

Full text available: pdf(71.02 KB) Additional Information: full citation, abstract, references

The paper contains a description of the Spoken Language Corpus of Swedish at the Department of Linguistics, Göteborg University (GSLC), and a summary of the various types of analysis and tools that have been developed for work on this corpus. Work on the corpus was started in the late 1970:s. It is incrementally growing and presently consists of 1.3 million words from about 25 different social activities. The corpus was initiated to meet a growing interest in naturalistic spoken language da ...

<u>Document detection</u>: Inquery system overview

John Broglio, James P. Callan, W. Bruce Croft

September 1993 Proceedings of a workshop on held at Fredericksburg, Virginia: September 19-23, 1993

Publisher: Association for Computational Linguistics

Additional Information: full citation, abstract, references Full text available: pdf(1.57 MB)

The TIPSTER project in the Information Retrieval Laboratory of the Computer Science Department, University of Massachusetts, Amherst (which includes MCC as a subcontractor), has focused on the following goals: • Improving the effectiveness of information retrieval techniques for large, full-text databases, Improving the effectiveness of routing techniques appropriate for long-term information needs, and. Demonstrating the effectiveness of these retrieval and routing techniques for ...

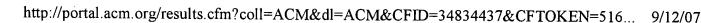
High-performance bilingual text alignment using statistical and dictionary information Masahiko Haruno, Takefumi Yamazaki

June 1996 Proceedings of the 34th annual meeting on Association for Computational Linquistics

Publisher: Association for Computational Linguistics

Full text available: mpdf(757.68 KB)





Publisher Site

Additional Information: full citation, abstract, references, citings

This paper describes an accurate and robust text alignment system for structurally different languages. Among structurally different languages such as Japanese and English, there is a limitation on the amount of word correspondences that can be statistically acquired. The proposed method makes use of two kinds of word correspondences in aligning bilingual texts. One is a bilingual dictionary of general use. The other is the word correspondences that are statistically acquired in the alignment pr ...

Noun-noun compound machine translation: a feasibility study on shallow processing Takaaki Tanaka, Timothy Baldwin



July 2003 Proceedings of the ACL 2003 workshop on Multiword expressions: analysis, acquisition and treatment - Volume 18

Publisher: Association for Computational Linguistics

Full text available: ndf(164.29 KB) Additional Information: full citation, abstract, references

The translation of compound nouns is a major issue in machine translation due to their frequency of occurrence and high productivity. Various shallow methods have been proposed to translate compound nouns, notable amongst which are memory-based machine translation and word-to-word compositional machine translation. This paper describes the results of a feasibility study on the ability of these methods to translate Japanese and English noun-noun compounds.

Technique for automatically correcting words in text



Karen Kukich

December 1992 ACM Computing Surveys (CSUR), Volume 24 Issue 4

Publisher: ACM Press

Full text available: mpdf(6.23 MB)

Additional Information: full citation, abstract, references, citings, index terms, review

Research aimed at correcting words in text has focused on three progressively more difficult problems:(1) nonword error detection; (2) isolated-word error correction; and (3) context-dependent work correction. In response to the first problem, efficient patternmatching and n-gram analysis techniques have been developed for detecting strings that do not appear in a given word list. In response to the second problem, a variety of general and application-specific spelling cor ...

Keywords: n-gram analysis, Optical Character Recognition (OCR), context-dependent spelling correction, grammar checking, natural-language-processing models, neural net classifiers, spell checking, spelling error detection, spelling error patterns, statisticallanguage models, word recognition and correction

10 A state of the art of Thai Language resources and Thai Language behavior analysis and modeling



Asanee Kawtrakul, Mukda Suktarachan, Patcharee Varasai, Hutchatai Chanlekha August 2002 Proceedings of the 3rd workshop on Asian language resources and international standardization - Volume 12 COLING '02

Publisher: Association for Computational Linguistics

Full text available: pdf(263.82 KB) Additional Information: full citation; abstract, references

As electronic communications is now increasing, the term Natural Language Processing should be considered in the broader aspect of Multi-Language processing system. Observation of the language behavior will provide a good basis for design of computational language model and also creating cost-effective solutions to the practical problems. In order to have a good language modeling, the language resources are necessary for the language behavior analysis. This paper intended to express what we have ...

11 N-gram-based Machine Translation

José B. Mariòo, Rafael E. Banchs, Josep M. Crego, Adrià de Gispert, Patrik Lambert, José A. R. Fonollosa, Marta R. Costa-jussà

December 2006 Computational Linguistics, Volume 32 Issue 4

Publisher: MIT Press

Full text available: pdf(291.73 KB) Additional Information: full citation, abstract

This article describes in detail an n-gram approach to statistical machine translation. This approach consists of a log-linear combination of a translation model based on n-grams of bilingual units, which are referred to as tuples, along with four specific feature functions. Translation performance, which happens to be in the state of the art, is demonstrated with Spanish-to-English and English-to-Spanish translations of the European Parliament Plenary Sessions (EPPS).

12 Computational methods ("paradigms"): High-probability syntactic links

Leonid Mitjushin

August 1992 Proceedings of the 14th conference on Computational linguistics -Volume 3

Publisher: Association for Computational Linguistics

Full text available: pdf(431.62 KB) Additional Information: full citation, references, citings

13 Searching the audio notebook: keyword search in recorded conversations Peng Yu, Kaijiang Chen, Lie Lu, Frank Seide



October 2005 Proceedings of the conference on Human Language Technology and **Empirical Methods in Natural Language Processing HLT '05**

Publisher: Association for Computational Linguistics

Full text available: pdf(452.53 KB) Additional Information: full citation, abstract, references

MIT's Audio Notebook added great value to the note-taking process by retaining audio recordings, e.g. during lectures or interviews. The key was to provide users ways to quickly and easily access portions of interest in a recording. Several non-speechrecognition based techniques were employed. In this paper we present a system to search directly the audio recordings by key phrases. We have identified the user requirements as accurate ranking of phrase matches, domain independence, and reasonabl ...

14 Phrase recognition and expansion for short, precision-biased queries based on a



query log

Erika F. de Lima, Jan O. Pedersen

August 1999 Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '99

Publisher: ACM Press

Full text available: pdf(189.83 KB) Additional Information: full citation, references, citings, index terms

15 Information extraction: The TIPSTER/SHOGUN project

Paul S. Jacobs, George Krupka, Lisa Rau, Michael L. Mauldin, Teruko Mitamura, Tsuyoshi Kitani, Ira Sider, Lois Childs

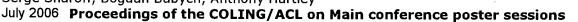
September 1993 Proceedings of a workshop on held at Fredericksburg, Virginia: September 19-23, 1993

Publisher: Association for Computational Linguistics

Full text available: pdf(1.11 MB) Additional Information: full citation, abstract, references

This paper presents an overview of the TIPSTER/SHOGUN project, the major results, and the SHOGUN data extraction system. TIPSTER/SHOGUN was a joint effort of GE Corporate Research and Development, Carnegie Mellon University, and Martin Marietta Management and Data Systems (formerly GE Aerospace), part of the ARPA TIPSTER Text program. Two of the main technical thrusts of the project were: (1) the development of a model of finitestate approximation, in which the accuracy of more detailed models ...

16 Using comparable corpora to solve problems difficult for human translators Serge Sharoff, Bogdan Babych, Anthony Hartley



Publisher: Association for Computational Linguistics

Full text available: pdf(151,23 KB) Additional Information: full citation, abstract, references

In this paper we present a tool that uses comparable corpora to find appropriate translation equivalents for expressions that are considered by translators as difficult. For a phrase in the source language the tool identifies a range of possible expressions used in similar contexts in target language corpora and presents them to the translator as a list of suggestions. In the paper we discuss the method and present results of human evaluation of the performance of the tool, which highlight it ...

17 Improving LSA-based summarization with anaphora resolution Josef Steinberger, Mijail A. Kabadjov, Massimo Poesio, Olivia Sanchez-Graillet October 2005 Proceedings of the conference on Human Language Technology and **Empirical Methods in Natural Language Processing HLT '05**

Publisher: Association for Computational Linguistics

Full text available: nodf(166.59 KB) Additional Information: full cliation, abstract, references

We propose an approach to summarization exploiting both lexical information and the output of an automatic anaphoric resolver, and using Singular Value Decomposition (SVD) to identify the main terms. We demonstrate that adding anaphoric information results in significant performance improvements over a previously developed system, in which only lexical terms are used as the input to SVD. However, we also show that how anaphoric information is used is crucial: whereas using this information to ad ...

18 Text Adaptation for Mobile Digital Teletext

Chengyuan Peng, Petri Vuorimaa

September 2004 Proceedings of the 2004 IEEE/WIC/ACM International Conference on Web Intelligence WI '04

Publisher: IEEE Computer Society

Full text available: pdf(111.90 KB)

Publisher Site

Additional Information: full citation, abstract

Small and varying screen sizes of mobile devices pose a big problem for digital Teletext service to display its content. It is difficult to display all the text information on a small screen, where page scroll or transparent page to live video is not practical. This paper presents an adaptive text extraction method which can automatically extract key information from original text and keep semantic meanings as close as possible. We combine both statistical methods and coarse coding algorithm fro ...

19 Combining trigram and Winnow in thai OCR error correction Surapant Meknavin, Boonserm Kijsirikul, Ananlada Chotimongkol, Cholwich Nuttee August 1998 Proceedings of the 36th annual meeting on Association for Computational Linguistics - Volume 2, Proceedings of the 17th international conference on Computational linguistics - Volume 2

Publisher: Association for Computational Linguistics

Full text available: pdf(592,57 KB) Additional Information: full citation, abstract, references, citings

For languages that have no explicit word boundary such as Thai, Chinese and Japanese, correcting words in text is harder than in English because of additional ambiguities in locating error words. The traditional method handles this by hypothesizing that every substrings in the input sentence could be error words and trying to correct all of them. In this paper, we propose the idea of reducing the scope of spelling correction by focusing only on dubious areas in the input sentence. Boundaries of ...

Text classification: Web-page classification through summarization



Dou Shen, Zheng Chen, Qiang Yang, Hua-Jun Zeng, Benyu Zhang, Yuchang Lu, Wei-Ying Ma July 2004 Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '04

Publisher: ACM Press

Full text available: pdf(225.80 KB)

Additional Information: full citation, abstract, references, citings, index terms, review

Web-page classification is much more difficult than pure-text classification due to a large variety of noisy information embedded in Web pages. In this paper, we propose a new Web-page classification algorithm based on Web summarization for improving the accuracy. We first give empirical evidence that ideal Web-page summaries generated by human editors can indeed improve the performance of Web-page classification algorithms. We then propose a new Web summarization-based classification algorithm ...

Keywords: content body, web page categorization, web page summarization

Results 1 - 20 of 42

Result page: 1 2 3 next

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2007 ACM, Inc.

<u>Terms of Usage Privacy Policy Code of Ethics Contact Us</u>

Useful downloads: Adobe Acrobat QuickTime Windows Media Player Real Player



Subscribe (Full Service) Register (Limited Service, Free) Login

Search: • The ACM Digital Library • The Guide

+"(word combination" or sentence or phrase or paragraph or v



THE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used: word combination or sentence or phrase or paragraph or words compute corpus segments Found 42 of

210,707

Sort results

relevance by

Results 21 - 40 of 42

Save results to a Binder 2 Search Tips

Try an Advanced Search Try this search in The ACM Guide

Display results

expanded form

Open results in a new window

Result page: previous 1 2 3

Relevance scale

Relevance

21 An empirical model of multiword expression decomposability

Timothy Baldwin, Colin Bannard, Takaaki Tanaka, Dominic Widdows

July 2003 Proceedings of the ACL 2003 workshop on Multiword expressions: analysis, acquisition and treatment - Volume 18

Publisher: Association for Computational Linguistics

Full text available: pdf(142.89 KB) Additional Information: full citation, abstract, references, citings

This paper presents a construction-inspecific model of multiword expression decomposability based on latent semantic analysis. We use latent semantic analysis to determine the similarity between a multiword expression and its constituent words, and claim that higher similarities indicate greater decomposability. We test the model over English noun-noun compounds and verb-particles, and evaluate its correlation with similarities and hyponymy values in WordNet. Based on mean hyponymy over partitio ...

22 Semantic role labeling of nominalized predicates in Chinese

Nianwen Xue

June 2006 Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of **Computational Linguistics**

Publisher: Association for Computational Linguistics

Full text available: ndf(105,35 KB) Additional Information: full citation, abstract, references

Recent work on semantic role labeling (SRL) has focused almost exclusively on the analysis of the predicate-argument structure of verbs, largely due to the lack of humanannotated resources for other types of predicates that can serve as training and test data for the semantic role labeling systems. However, it is well-known that verbs are not the only type of predicates that can take arguments. Most notably, nouns that are nominalized forms of verbs and relational nouns generally are also consi ...

23 tRuEcasing

Lucian Vlad Lita, Abe Ittycheriah, Salim Roukos, Nanda Kambhatla

July 2003 Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1 ACL '03

Publisher: Association for Computational Linguistics

Full text available: pdf(173.84 KB) Additional Information: full citation, abstract, references, citings

Truecasing is the process of restoring case information to badly-cased or non-cased text. This paper explores truecasing issues and proposes a statistical, language modeling based

http://portal.acm.org/results.cfm?query=%2B%22%28word%20combination%22%20or%20s... 9/12/07

truecaser which achieves an accuracy of ~98% on news articles. Task based evaluation shows a 26% F-measure improvement in named entity recognition when using truecasing. In the context of automatic content extraction, mention detection on automatic speech recognition text is also improved by a factor of 8. Truec ...

A method for open-vocabulary speech-driven text retrieval

Atsushi Fujii, Katunobu Itou, Tetsuya Ishikawa

July 2002 Proceedings of the ACL-02 conference on Empirical methods in natural language processing - Volume 10 EMNLP '02

Publisher: Association for Computational Linguistics

Full text available: ndf(107.74 KB) Additional Information: full citation, abstract, references, citings

While recent retrieval techniques do not limit the number of index terms, out-ofvocabulary (OOV) words are crucial in speech recognition. Aiming at retrieving information with spoken queries, we fill the gap between speech recognition and text retrieval in terms of the vocabulary size. Given a spoken query, we generate a transcription and detect OOV words through speech recognition. We then correspond detected OOV words to terms indexed in a target collection to complete the transcription, and ...

25 Summarization: Web-page summarization using clickthrough data

Jian-Tao Sun, Dou Shen, Hua-Jun Zeng, Qiang Yang, Yuchang Lu, Zheng Chen August 2005 Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '05

Publisher: ACM Press

Additional Information: full citation, abstract, references, citings, index Full text available: pdf(157.70 KB) terms

Most previous Web-page summarization methods treat a Web page as plain text. However, such methods fail to uncover the full knowledge associated with a Web page needed in building a high-quality summary, because many of these methods do not consider the hidden relationships in the Web. Uncovering the hidden knowledge is important in building good Web-page summarizers. In this paper, we extract the extra knowledge from the clickthrough data of a Web search engine to improve Web-page summarization ...

Keywords: clickthrough data, generic web-page summarization, latent semantic analysis, thematic lexicon

26 CoNLL-2000 papers: Increasing our ignorance of language: identifying language structure in an unknown 'signal'

John Elliott, Eric Atwell, Bill Whyte

September 2000 Proceedings of the 2nd workshop on Learning language in logic and the 4th conference on Computational natural language learning -Volume 7

Publisher: Association for Computational Linguistics

Full text available: pdf(588.40 KB) Additional Information: full citation, abstract, references

This paper describes algorithms and software developed to characterise and detect generic intelligent language-like features in an input signal, using natural language learning techniques: looking for characteristic statistical "language-signatures" in test corpora. As a first step towards such species-independent language-detection, we present a suite of programs to analyse digital representations of a range of data, and use the results to extrapolate whether or not there are language-like stru ...

27 NEXUS: a linguistic technique for procoordination

R. A. Benson

September 1969 Proceedings of the 1969 conference on Computational linguistics

Publisher: Association for Computational Linguistics

Full text available: pdf(1.31 MB) Additional Information: full citation, abstract, references

A method for automatically precoordinating index terms was devised to form combinations of terms which are stored as subject headings. A computer program accepts lists of autoindexed terms and by applying linguistic and sequence rules combines appropriate terms, thereby effecting improved searchability of an information storage and retrieval system.A serious failing exists in many indexing systems in that index terms authorized for use are too general for use by technically-knowledgeable search ...

28 Fast and quasi-natural language search for gigabytes of Chinese texts

Lee-Feng Chien

July 1995 Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '95

Publisher: ACM Press

Full text available: pdf(820.56 KB) Additional Information: full citation, references, citings, index terms

Book reviews

Karen Sutherland

March 1999 intelligence, Volume 10 Issue 1

Publisher: ACM Press

Full text available: pdf(450.61 KB) html(35.73 KB)

Additional Information: full citation, references, index terms

30 Poster Sessions: A gradual refinement model for a robust thai morphological analyzer



Kawtrakul Asanee, Thumkanon Chalatip, Jamjanya Thitima, Muangyunnan Parinee, Poolwan Kritsada, Inagaki Yasuyoshi

August 1996 Proceedings of the 16th conference on Computational linguistics -Volume 2

Publisher: Association for Computational Linguistics

Full text available: pdf(360.47 KB) Additional Information: full citation, abstract, references

This work attempts to provide a robust Thai morphological analyzer which can automatically assign the correct part-of-speech tag to the correct word with time and space efficiency. Instead of using a corpus based approach with requires a large amount of training data and validation data, a new simple hybrid technique which incorporates heuristic, syntactic and semantic knowledge is proposed. To implement this technique, a three-stage approach is adopted to the gradual refinement module. It consi ...

31 Special issue on Machine learning methods for text and images: Word sequence kernels



Nicola Cancedda, Eric Gaussier, Cyril Goutte, Jean Michel Renders March 2003 The Journal of Machine Learning Research, Volume 3

Publisher: MIT Press

Full text available: pdf(208.06 KB) Additional Information: full citation, abstract, citings, index terms

We address the problem of categorising documents using kernel-based methods such as Support Vector Machines. Since the work of Joachims (1998), there is ample experimental evidence that SVM using the standard word frequencies as features yield state-of-the-art performance on a number of benchmark problems. Recently, Lodhi et al. (2002) proposed the use of string kernels, a novel way of computing document similarity based of matching non-consecutive subsequences of characters. In this arti ...

32 Creating segmented databases from free text for text retrieval

Lisa F. Rau, Paul S. Jacobs

September 1991 Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '91

Publisher: ACM Press

Full text available: pdf(904.82 KB) Additional Information: full citation, references, citings, index terms

33 Statistical language modeling: Hypothesizing word association from untagged text Tomovoshi Matsukawa

March 1993 Proceedings of the workshop on Human Language Technology HLT '93

Publisher: Association for Computational Linguistics

Full text available: pdf(498.43 KB) Additional Information: full citation, abstract, references, citings

This paper reports a new method for suggesting word associations, based on a greedy algorithm that employs Chi-square statistics on joint frequencies of pairs of word groups compared against chance co-occurrence. The benefits of this new approach are: 1) we can consider even low frequency words and word pairs, and 2) word groups and word associations can be automatically generated. The method provided 87% accuracy in hypothesizing word associations for unobserved combinations of words in Japanes ...

34 Semantics: Generalizing automatically generated selectional patterns Ralph Grishman, John Sterling



August 1994 Proceedings of the 15th conference on Computational linguistics -Volume 2

Publisher: Association for Computational Linguistics

Full text available: pdf(461.98 KB) Additional Information: full citation, abstract, references, citings

Frequency information on co-occurrence patterns can be automatically collected from a syntactically analyzed corpus; this information can then serve as the basis for selectional constraints when analyzing new text from the same domain. This information, however, is necessarily incomplete. We report on measurements of the degree of selectional coverage obtained with different sizes of corpora. We then describe a technique for using the corpus to identify selectionally similar terms, and for using ...

Abbreviation Expansion in Schema Matching and Web Integration



L. Ratinov, E. Gudes

September 2004 Proceedings of the 2004 IEEE/WIC/ACM International Conference on Web Intelligence WI '04

Publisher: IEEE Computer Society

Full text available: pdf(112.64 KB) Publisher Site

Additional Information: full citation, abstract

Schema matching is a problem of finding correspondences, particularly equivalence relationships across schemas. The problem has a particular significance in integrating web repositories, as distributed databases over the web becomes increasingly popular. Most of the existing prototypes use schema level lexical information for schema matching. However, most of them perform rather poorly on real-world problems due to the

abundance of abbreviations in real-world schemas. For example, none of the le ...

36 Co-ordinative ellipsis in Russian texts: problems of description and restoration Igor A. Bolshakov



August 1988 Proceedings of the 12th conference on Computational linguistics -Volume 1

Publisher:	Association	for (Computational L	inquietice
ruviisiiei.	ASSOCIATION	IUI V	Julibulalional L	unuuisiics

Full text available: pdf(477.65 KB) Additional Information: full citation, abstract, references

Russian elliptic constructions are examined from the point of view of syntactic analysis. Reciprocal elements in a co-ordinative elliptic sentence are exposed and possible types of their similarity are explored. Linear formulae of ellipsis for most textual cases are constructed and statistics of their use is discussed. As a result the main steps of ellipsis restoration algorithm are outlined.

37 Automatic learning of language model structure

Kevin Duh, Katrin Kirchhoff

August 2004 Proceedings of the 20th international conference on Computational Linguistics COLING '04

Publisher: Association for Computational Linguistics

Full text available: pdf(167.26 KB) Additional Information: full citation, abstract, references

Statistical language modeling remains a challenging task, in particular for morphologically rich languages. Recently, new approaches based on factored language models have been developed to address this problem. These models provide principled ways of including additional conditioning variables other than the preceding words, such as morphological or syntactic features. However, the number of possible choices for model parameters creates a large space of models that cannot be searched exhaustive ...

38 Viewing stemming as recall enhancement

Wessel Kraaij, Renée Pohlmann

August 1996 Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '96

Publisher: ACM Press

Full text available: pdf(928.26 KB) Additional Information: full citation, references, citings, index terms

Extracting clauses for spoken language understanding in conversational systems

Narendra K. Gupta, Srinivas Bangalore

July 2002 Proceedings of the ACL-02 conference on Empirical methods in natural language processing - Volume 10 EMNLP '02

Publisher: Association for Computational Linguistics

Full text available: pdf(228.90 KB) Additional Information: full citation, abstract, references

Spontaneous human utterances in the context of human-human and human-machine dialogs are rampant with dysfluencies, and speech repairs. Furthermore, when recognized using a speech recognizer, these utterances produce a sequence of words with no identification of clausal units. Such long strings of words combined with speech errors pose a difficult problem for spoken language parsing and understanding. In this paper, we address the issue of editing speech repairs as well as segmenting user uttera ...

40 Knowledge-free induction of inflectional morphologies

Patrick Schone, Daniel Jurafsky

June 2001 Second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies 2001 NAACL '01

Publisher: Association for Computational Linguistics

Full text available: pdf(261.90 KB) Additional Information: full citation, abstract, references, citings

We propose an algorithm to automatically induce the morphology of inflectional languages using only text corpora and no human input. Our algorithm combines cues from orthography, semantics, and syntactic distributions to induce morphological relationships in German, Dutch, and English. Using CELEX as a gold standard for evaluation, we show our algorithm to be an improvement over any knowledge-free algorithm yet proposed.

Results (page 2): +"(word combination" or sentence or phrase or paragraph or words)" +c... Page 6 of 6

Results 21 - 40 of 42

Result page: previous 1 2 3 next

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2007 ACM, Inc.

<u>Terms of Usage Privacy Policy Code of Ethics Contact Us</u>

Useful downloads: Adobe Acrobat QuickTime Windows Media Player Real Player



Subscribe (Full Service) Register (Limited Service, Free) Login

Search: The ACM Digital Library The Guide

+"(word combination" or sentence or phrase or paragraph or v

THE ACM DIG TAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used: word combination or sentence or phrase or paragraph or words compute corpus segments Found 42 of

210,707

Sort results by

relevance

Save results to a Binder 2 Search Tips

Try an Advanced Search Try this search in The ACM Guide

Display results

expanded form

Open results in a new

window

Results 41 - 42 of 42

Result page: previous 1 2 3

Relevance scale ...

41 Information retrieval: Speech-based retrieval using semantic co-occurrence filtering Julian Kupiec, Don Kimber, Vijay Balasubramanian

March 1994 Proceedings of the workshop on Human Language Technology HLT '94

Publisher: Association for Computational Linguistics

Full text available: pdf(440.55 KB) Additional Information: full citation, abstract, references, citings

In this paper we demonstrate that speech recognition can be effectively applied to information retrieval (IR) applications. Our system exploits the fact that the intended words of a spoken query tend to co-occur in text documents in close proximity whereas word combinations that are the result of recognition errors are usually not semantically correlated and thus do not appear together. Termed "Semantic Co-occurrence Filtering" this enables the system to simultaneously disambiguate word hypothes ...

42 A stochastic process for word frequency distributions

Harald Baayen

June 1991 Proceedings of the 29th annual meeting on Association for Computational Linguistics

Publisher: Association for Computational Linguistics

Full text available: pdf(673.84 KB)

Publisher Site

Additional Information: full citation, abstract, references, citings

A stochastic model based on insights of Mandelbrot (1953) and Simon (1955) is discussed against the background of new criteria of adequacy that have become available recently as a result of studies of the similarity relations between words as found in large computerized text corpora.

Results 41 - 42 of 42

Result page: previous 1 2 3

The ACM Portal is published by the Association for Computing Machinery. Copyright @ 2007 ACM, Inc. Terms of Usage Privacy Policy Code of Ethics Contact Us

Useful downloads: Adobe Acrobat QuickTime Windows Media Player Real Player



Welcome United States Patent and Trademark Office

	Results	'nΡ	Searc
--	---------	-----	-------

Search Resu	ılts	0000	erowse search ieee xplore guide
Your search	((word <in>metadata)<and>(comatched 46 of 1641691 documents of 100 results are displayed, 25 to a</and></in>	s . ,	>metadata)) <and> (corpus<"</and>
» Search Opt	dons	Modify S	Search
View Session	n.History	((word<	in>metadata) <and> (combination<in>metadata))<and> (corpus<in>meta</in></and></in></and>
New Search	•		eck to search only within this results set
		Display I	
» Key		Mainer 1	Gitation & Abstract
ieee jnl	IEEE Journal or Magazine	← view	Selected items Select All Deselect All
iet jal	IET Journal or Magazine		•
ieee onf	IEEE Conference Proceeding		Combining Multi-knowledge for Chinese Word Segmentation Disambiguation Ying Qin; Suxiang Zhang; Xiaojie Wang;
BET ONF	IET Conference Proceeding		Intelligent Systems Design and Applications, 2006, ISDA '06, Sixth International Conference on
ieee std	IEEE Standard		Volume 1, Oct. 2006 Page(s):551 - 556 Digital Object Identifier 10.1109/ISDA.2006.124
			AbstractPlus Full Text: PDF(324 KB) IEEE CNF Rights and Permissions
	·		2. Integrating tonal information into Mandarin name recognition with different strategles Dong-sheng Luo; Xiang Xie; Jing-ming Kuang; Chinese Spoken Language Processing, 2004 International Symposium on 15-18 Dec. 2004 Page(s):265 - 268 Digital Object Identifier 10.1109/CHINSL.2004.1409637
	•	. •	AbstractPlus Full Text: PDF(609 KB) INNECTIONS Rights and Permissions
	. •	□ 3	3. A shape based post processor for Gurmukhi OCR Lehal, G.S.; Singh, C.; Lehal, R.; Document Analysis and Recognition. 2001. Proceedings. Sixth International Conference on 10-13 Sept. 2001 Page(s):1105 - 1109 Digital Object Identifier 10.1109/ICDAR.2001.953957 AbstractPlus Full Text: PDE(424 KB) ISSE CRE
		·····	Rights and Permissions
		1 4	Acoustics. Speech, and Signal Processing, 1999. ICASSP '99. Proceedings, 1999. IEEE Internation Volume 1, 15-19 March 1999. Page(s):529 - 532 vol.1 Digital Object Identifier 10.1109/ICASSP.1999.758179 AbstractPlus Full Text: PDF(328 KB) IEEE CNF Rights and Permissions
		5	5. Using boosting to improve a hybrid HMM/neural network speech recognizer Schwenk, H.; Accustics, Speech, and Signal Processing, 1999. ICASSP '99. Proceedings, 1999 IEEE Internation Volume 2, 15-19 March 1999 Page(s):1009 - 1012 vol.2 Digital Object Identified 10,4400/04000 1009 - 1

AbstractPlus | Full Text: PDE(332 KB) ISSE ONF Rights and Permissions 6. On the dynamic adaptation of stochastic language models Kneser, R.; Steinbiss, V.; Acoustics, Speech, and Signal Processing, 1993, ICASSP-93, 1993 IEEE International Conference Volume 2, 27-30 April 1993 Page(s):586 - 589 vol.2 Digital Object Identifier 10.1109/ICASSP.1993.319375 AbstractPlus | Full Text: PDE(280 KB) #EEE CNF Rights and Permissions 7. Dialect/Accent Classification Using Unrestricted Audio Huang, R.; Hansen, J. H. L.; Angkititrakul, P.; Audio Speech and Language Processing IEEE Transactions on [see also Speech and Audio Proc Transactions on Volume 15, issue 2, Feb. 2007 Page(s):453 - 464 Digital Object Identifier 10.1109/TASL.2006.881695 AbstractPlus | Full Text: PDF(620 KB) (SIBIL JN). Rights and Permissions 8. Joint Optimization of Word Alignment and Epenthesis Generation for Chinese to Taiwanese Yu-Hsien Chiu; Chung-Hsien Wu; Hung-Yu Su; Chih-Jen Cheng; Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 29, issue 1, Jan. 2007 Page(s):28 - 39 Digital Object Identifier 10.1109/TPAMI.2007.250597 AbstractPlus | Full Text: PDE(3770 KB) IIIIIII JAN. Rights and Permissions 9. Automatic indexing of lecture presentations using unsupervised learning of presumed disco Kawahara, T.; Hasegawa, M.; Shitaoka, K.; Kitade, T.; Nanjo, H.; Speech and Audio Processing, IEEE Transactions on Volume 12, Issue 4, July 2004 Page(s):409 - 419 Digital Object Identifier 10.1109/TSA.2004.828701 AbstractPlus | References | Full Text: PDF(360 KB) | IEEE JNIL Rights and Permissions 10. Unsupervised training of acoustic models for large vocabulary continuous speech recogniti Wessel, F.; Ney, H.; Speech and Audio Processing, IEEE Transactions on Volume 13, issue 1, Jan. 2005 Page(s):23 - 31 Digital Object Identifier 10.1109/TSA.2004.838537 AbstractPlus | References | Full Text: PDF(488 KB) | ISSE JNL Rights and Permissions 11. Topic and Stylistic Adaptation for Speech Summarisation Chatain, P.; Whittaker, E.W.D.; Mrozinski, J.A.; Furui., S.; Acquistics, Speech and Signal Processing, 2006, ICASSP 2006 Proceedings, 2006 IEEE Internation Volume 1, 14-19 May 2006 Page(s): I - I Digital Object Identifier 10.1109/ICASSP.2006.1660186 AbstractPlus | Full Text: PDF(169 KB) IEEE ONF Rights and Permissions 12. Dynamic Bayesian Networks and Discriminative Classifiers for Multi-Stage Semantic Interpr Acoustics, Speech and Signal Processing, 2007, ICASSP 2007, IEEE International Conference on Volume 4, 15-20 April 2007 Page(s):IV-13 - IV-16 Digital Object Identifier 10.1109/ICASSP.2007.367151 AbstractPlus | Full Text: PDF(105 KB) 1888 CRF Rights and Permissions

13. Word-Conditioned Phone N-Grams for Speaker Recognition Lei, H.; Mirghafori, N.; Acoustics, Speech and Signal Processing, 2007, iCASSP 2007, IEEE International Conference on Volume 4, 15-20 April 2007 Page(s):IV-253 - IV-256 Digital Object Identifier 10.1109/ICASSP.2007.366897
AbstractPlus Full Text: PDF(4537 KB) ISSE CRF Rights and Permissions
14. Identifying Nocuous Ambiguitles in Natural Language Requirements Requirements Engineering Conference, 2006. RE 2006. 14th IEEE International 2006 Page(s):56 - 65 Digital Object Identifier 10.1109/RE.2006.31 AbstractPlus Full Text: PDF(441 KB) 接触 CNF Rights and Permissions
15. Computational Approach for Processing of Control Engineer Text: Applications for Corpus I Vitayapirak, J.; Ratiroch-anant, P.; Oybernetics and Intelligent Systems, 2006 IEEE Conference on June 2006 Page(s):1 - 6 Digital Object Identifier 10.1109/ICCIS.2006.252266 AbstractPlus Full Text: PDE(279 KB) IEEE CNF Rights and Permissions
16. Combining statistical similarity measures for automatic Induction of semantic classes Pangos, A.; Iosif, E.; Potamianos, A.; Fosler-Lussier, E.; Automatic Speech Recognition and Understanding, 2005 IEEE Workshop on 27 Nov1 Dec. 2005 Page(s):278 - 283 Digital Object Identifier 10.1109/ASRU.2005.1566510 AbstractPlus Full Text: PDE(195 KB) \$\frac{1811111}{1811111} \text{CNF}\$ Rights and Permissions
17. Sentence Structure for Dialog Act Recognition in Czech Kral, P.; Cerisara, C.; Kleckova, J.; Pavelka, T.; Information and Communication Technologies. 2006. ICTTA '06. 2nd Volume 1, 24-28 April 2006 Page(s):1214 - 1218 AbstractPlus Full Text: PDF(1000 KB)
18. Hierarchical iterative and self-supervised method for concept-word acquisition from large-sing Guogang Tian; Cungen Cao; Natural Language Processing and Knowledge Engineering, 2005, IFEE NLP-KE '05, Proceedings in International Conference on 30 Oct1 Nov. 2005 Page(s):312 - 317 Digital Object Identifier 10.1109/NLPKE.2005.1598754 AbstractPlus Full Text: PDE(1032 KB) ISSE CNF Rights and Permissions
19. Acoustic model combination for recognition of speech in multiple languages using support Gangashetty, S.V.; Sekhar, C.C.; Yegnanarayana, B.; Neural Networks, 2004. Proceedings, 2004: IEEE International Joint Conference on Volume 4, 25-29 July 2004 Page(s):3065 - 3069 vol.4 AbstractPlus Full Text: PDE(625 KB) IEEE CNF Rights and Permissions
20. Automatic indexing of key sentences for lecture archives using statistics of presumed disconance, Nanjo, H.; Kitade, T.; Kawahara, T.; Acoustics, Speech, and Signal Processing, 2004, Proceedings, (ICASSP '04), IEEE International Communications of the processing of the processin

Digital Object Identifier 10.1109/ICASSP.2004.1326019 AbstractPlus | Full Text: PDF(243 KB) ISEE CNF Rights and Permissions 21. Warping and scaling of the minimum variance distortionless response Wolfel, M.; McDonough, J.; Waibel, A.; Automatic Speech Recognition and Understanding, 2003, ASRU '03, 2003 IEEE Workshop on 30 Nov.-3 Dec. 2003 Page(s):387 - 392 Digital Object Identifier 10.1109/ASRU.2003.1318472 AbstractPlus | Full Text: PDF(446 KB) (\$1838 CINF) Rights and Permissions 22. Automatic segmentation of Chinese chunks using a neural network Rongbo Wang; Zheru Chi; Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference of Volume 1, 14-17 Dec. 2003 Page(s):96 - 99 Vol.1 Digital Object Identifier 10.1109/ICNNSP.2003.1279221 AbstractPlus | Full Text: PDF(309 KB) | III EE CMF Rights and Permissions 23. Integrated Chinese word segmentation and part-of-speech tagging based on the divide-and-Sun Maosong; Xu Dongliang; Tsou, B.K.; Natural Language Processing and Knowledge Engineering, 2003. Proceedings, 2003 International 26-29 Oct. 2003 Page(s):610 - 615 Digital Object Identifier 10.1109/NLPKE.2003.1275978 AbstractPlus | Full Text: PDF(397 KB) IEEE CNF Rights and Permissions 24. Parsing N-best lists of handwritten sentences Zimmermann, M.; Chappelier, J.-C.; Bunke, H.; Document Analysis and Recognition, 2003. Proceedings. Seventh International Conference on 3-6 Aug. 2003 Page(s):572 - 576 vol.1 Digital Object Identifier 10.1109/ICDAR.2003.1227729 AbstractPlus | Full Text: PDF(282 KB) IEEE CNF Rights and Permissions 25. Improving the performance of an LVCSR system through ensembles of acoustic models П Rong Zhang; Rudnicky, A.I.; Acoustics, Speech, and Signal Processing, 2003, Proceedings, (ICASSP '03), 2003 IEEE International Control of the Control of t Volume 1, 6-10 April 2003 Page(s):1-876 - I-879 vol.1 AbstractPlus | Full Text: PDF(343 KB) IEEE CRF Rights and Permissions

Volume 1, 17-21 May 2004 Page(s):I - 449-52 vol.1

ilinspec

Help Contact Us Privac

& Copyright 2006 (S



Home $\{\text{ Logic }\}$ -Logicut $\{\text{ Access Information }\}$ -Alerts $\{\text{ Purchase History }\}$, or

Welcome United States Patent and Trademark Office

Sea Sea	ırch	Resul	te

Search Resu	ılts	erowse search ieee xplore guide
Your search	matched 46 of 1641691 documents.	mbination <in>metadata))<and> (corpus<"</and></in>
» Search Opt	ions	Dr. diffe Coarse
View Section	a History	Modify Search ((word <in>metadata) <and>(combination<in>metadata))<and>(corpus<in>metadata)</in></and></in></and></in>
View Session	THISIOIA.	((Moto sin metadata) samo (comunicatoris in metadata) samo (corpus sin metadata)
New Search		Check to search only within this results set
ь Кеу		Display Format;
BESE JNG.	IEEE Journal or Magazine	view selected items Select All Deselect All
BET JNL	IET Journal or Magazine	Selectali Deselectali
IEEE CNF	IEEE Conference Proceeding	26. Prosodic knowledge sources for automatic speech recognition
HIT CNF	IET Conference Proceeding	Vergyri, D.; Stolcke, A.; Gadde, V.R.R.; Ferrer, L.; Shriberg, E.;
IEEE STD	IEEE Standard	Acoustics, Speech, and Signal Processing, 2003, Proceedings, (ICASSP '03), 2003 IEEE Internation Volume 1, 6-10 April 2003 Page(s):1-208 - 1-211 vol.1
		AbstractPlus Full Text: PDE(389 KB) ISSE CNF Rights and Permissions
		27. Minimum phone error and I-smoothing for improved discriminative training Povey, D.; Woodland, P.C.; Acoustics, Speech, and Signal Processing, 2002, Proceedings, (ICASSP '02), IEEE International C
		Volume 1, 2002 Page(s):I-105 - I-108 vol.1 Digital Object Identifier 10.1109/ICASSP.2002.1005687
		AbstractPlus Full Text: <u>PDF(</u> 365 KB) #표표도 CNF Rights and <u>Permissions</u>
		28. Robust dialogue-state dependent language modeling using leaving-one-out Wessel, F.; Baader, A.; Acoustics. Speech, and Signal Processing, 1999. ICASSP '99. Proceedings, 1999 IEEE Internation Volume 2, 15-19 March 1999 Page(s):741 - 744 vol.2
٠,		Digital Object Identifier 10.1109/ICASSP.1999.759773 AbstractPlus Full Text: PDF(364 KB) → ISES CNF Rights and Permissions
		29. Issues in text-to-speech synthesis Macchi, M.; Intelligence and Systems 1998. Proceedings, IEEE International Joint Symposia on 21-23 May 1998 Page(s):318 - 325 Digital Object Identifier 10.1109/IJSIS.1998.685467
		AbstractPlus Full Text: <u>PDF(</u> 96 KB) 超距 CNF Rights and Permissions
		30. Improved lexical tree search for large vocabulary speech recognition Ortmanns, S.; Eiden, A.; Ney, H.; Acoustics. Speech, and Signal Processing, 1998, ICASSP '98, Proceedings of the 1998 IEEE Inter
		<u>on</u> Volume 2, 12-15 May 1998 Page(s):817 - 820 vol.2 Digital Object Identifier 10.1109/ICASSP.1998.675390

AbstractPlus | Full Text: PDE(388 KB) ISSE ONF Rights and Permissions 31. Initial language models for spoken dialogue systems Kellner, A.; Acoustics, Speech, and Signal Processing, 1998, ICASSP '98, Proceedings of the 1998 IEEE Inter-Volume 1, 12-15 May 1998 Page(s):185 - 188 vol.1 Digital Object Identifier 10.1109/ICASSP.1998.674398 AbstractPlus | Full Text: PDF(360 KB) INEE CNF Rights and Permissions 32. Look-ahead techniques for fast beam search Ortmanns, S.; Eiden, A.; Ney, H.; Coenen, N.; Acoustics, Speech, and Signal Processing, 1997, ICASSP-97, 1997 IEEE International Conference Volume 3, 21-24 April 1997 Page(s):1783 - 1786 vol.3 Digital Object Identifier 10.1109/ICASSP.1997.598876 AbstractPlus | Full Text: PDF (376 KB) | IEIEE CNF Rights and Permissions 33. Domain word translation by space-frequency analysis of context length histograms Fung. P.: Acoustics, Speech, and Signal Processing, 1996, ICASSP-96, Conference Proceedings, 1996 IEE Conference on Volume 1, 7-10 May 1996 Page(s):184 - 187 vol. 1 Digital Object Identifier 10.1109/ICASSP.1996.540321 AbstractPlus | Full Text: PDF(356 KB) REEE CNF Rights and Permissions 34. Combination of word-based and category-based language models ____ Niesler, T.R.; Woodland, P.C.; Spoken Language, 1996, ICSLP 96, Proceedings, Fourth International Conference on Volume 1, 3-6 Oct. 1996 Page(s):220 - 223 vol.1 Digital Object Identifier 10.1109/ICSLP.1996,607081 AbstractPlus | Full Text: PDF(284 KB) | IEEE CNF Rights and Permissions 35. Estimating Markov model structures _ Brants, T.: Spoken Language, 1996. ICSLP 96. Proceedings. Fourth International Conference on Volume 2, 3-6 Oct. 1996 Page(s):893 - 896 vol.2 Digital Object Identifier 10.1109/ICSLP.1996.607745 AbstractPlus | Full Text: PDE(340 KB) IEEE CNF Rights and Permissions 36. The design of a neural data-oriented parsing (DOP) system Scholtes, J.C.; Bloembergen, S.; Neural Networks, 1992, IJCNN, International Joint Conference on Volume 2, 7-11 June 1992 Page(s):69 - 74 vol.2 Digital Object Identifier 10.1109/IJCNN.1992.226982 AbstractPlus | Full Text: PDF(472 KB) ISEE CNF Rights and Permissions 37. The estimation of powerful language models from small and large corpora Placeway, P.; Schwartz, R.; Fung, P.; Nguyen, L.; Acoustics, Speech, and Signal Processing, 1993, ICASSP-93, 1993 IEEE International Conference Volume 2, 27-30 April 1993 Page(s):33 - 36 vol.2

Rights and Permissions

	38. Advances in transcription of broadcast news and conversational telephone speech within the BBN/LIMSI system
	Matsoukas, S.; Gauvain, JL.; Adda, G.; Colthurst, T.; Chia-Lin Kao; Kimball, O.; Lamel, L.; Lefevre
	Makhoul, J.; Nguyen, L.; Prasad, R.; Schwartz, R.; Schwenk, H.; Bing Xiang;
	Audio, Speech and Language Processing, IEEE Transactions on [see also Speech and Audio Proc
	Transactions on]
	Volume 14, <u>Issue 5</u> , Sept. 2006 Page(s):1541 - 1556
	Digital Object Identifier 10.1109/TASL.2006.878257
	AbstractPlus Full Text: PDF (1416 KB) ###################################
	·
	39. Segmental minimum Bayes-risk decoding for automatic speech recognition
	Goel, V.; Kumar, S.; Byrne, W.;
	Speech and Audio Processing. IEEE Transactions on
	Volume 12, <u>issue 3</u> , May 2004 Page(s):234 - 249
	Digital Object Identifier 10.1109/TSA.2004.825678
	AbstractPlus References Full Text: PDF(592 KB) : IEEE JML.
	Rights and Permissions
3000	40. Word sense disambiguation combining conceptual distance, frequency and gloss
	Rosso, P.; Masulli, F.; Buscaldi, D.;
	Natural Language Processing and Knowledge Engineering, 2003, Proceedings, 2003 International
	26-29 Oct. 2003 Page(s):120 - 125
	Digital Object Identifier 10.1109/NLPKE.2003.1275880
	AbstractPlus Full Text: PDF(389 KB) RESE CNF
	Rights and Permissions
	41. A Bayesian framework for fusing multiple word knowledge models in videotext recognition
	DongQuin Zhang; Shih-Fu Chang;
	Computer Vision and Pattern Recognition, 2003, Proceedings, 2003 IEEE Computer Society Confe
	Volume 2, 18-20 June 2003 Page(s):II-528 - II-533 vol.2
	AbstractPlus Full Text: PDF (266 KB) IEEE CNF
	Rights and Permissions
	42. Doing away with the Viterbi approximation
	Demuynck, K.; Van Compernolle, D.; Wambacq, P.;
	Acoustics, Speech, and Signal Processing, 2002, Proceedings, (ICASSP '02), IEEE international C
J	Volume 1, 2002 Page(s):I-717 - I-720 vol.1 Digital Object Identifier 10.1109/ICASSP.2002.1005840
	·
	AbstractPlus Full Text: PDF(379 KB) INEE CNF
	Rights and Permissions
	
	43. Text compression by context tree weighting
•	Aberg, J.; Shtarkov, Yu.M.;
	Data Compression Conference, 1997, DCC '97, Proceedings 25-27 March 1997 Page(s):377 - 386
	Digital Object Identifier 10.1109/DCC.1997.582062
	AbstractPlus Full Text: <u>PDF(536 KB)</u> ISSE ONF Rights and Permissions
بسبو	44 Predicting the out-of-vocabulary rate and the remised was believed.
1	44. Predicting the out-of-vocabulary rate and the required vocabulary size for speech processin Muller, J.; Stahl, H.; Lang, M.;
	Spoken Language, 1996, ICSLP 96, Proceedings, Fourth International Conference on
	Volume 3, 3-6 Oct. 1996 Page(s):1922 - 1925 vol.3
	Digital Object Identifier 10.1109/ICSLP.1996.608010
	AbstractPlus Full Text: PDF(288 KB) ISSE CNF
	Rights and Permissions

45. Word predictability after hesitations: a corpus-based study

Shriberg, E.; Stolcke, A.;

Spoken Language, 1996. ICSLP 96. Proceedings. Fourth International Conference on

Volume 3, 3-6 Oct. 1996 Page(s):1868 - 1871 vol.3

Digital Object Identifier 10.1109/ICSLP.1996.607996

AbstractPlus | Full Text: PDE(364 KB) ISSE CNF

Rights and Permissions

46. HAL-Based Evolutionary Inference for Pattern Induction From Psychiatry Web Resources

Yu L.-C. ; Wu C.-H. ; Yeh J.-F. ; Jang F.-L. ;

IEEE Transactions on Evolutionary Computation : Accepted for future publication

Volume PP, Issue 99, 2007 Page(s):1 - 1

Digital Object Identifier 10.1109/TEVC.2007.895270

AbstractPlus | Full Text: PDF(784 KB) IMME SINIL

Contact Us Privac

lesseed by



Home | Logout | Access Information | Alerts | Purchase History | " C.

Welcome United States Patent and Trademark Office

BROWSE

SEARCH

IEEE XPLORE GUIDE

∜ View Search Results | ∜ Previous Article | Next Article >

∭ e-mail

Access this document



Full Text: PDF (284 KB)

Download this citation

Choose Citation & Abstract

Download ASCII Text

Rights and Permissions

» Learn More

» Learn More

Combination of word-based and category-based language models

Niesler, T.R. Woodland, P.C.

Dept. of Eng., Cambridge Univ., UK;

This paper appears in: Spoken Language, 1996, ICSLP 96, Proceedings, Fourth International Conference on

Publication Date: 3-6 Oct. 1996

Volume: 1

On page(s): 220 - 223 vol.1 Number of Pages: 4 vol. 2522

Meeting Date: 10/03/1996 - 10/06/1996

Location: Philadelphia, PA

INSPEC Accession Number:5671314

Digital Object Identifier: 10.1109/ICSLP.1996.607081

Posted online: 2002-08-06 20:31:29.0

Abstract

A language model combining word based and category based n grams within a backoff framework is presented. Word capture sequential relations between particular words, while the category model, which is based on part of speech class ambiguous category membership, is able to generalise to unseen word sequences and therefore appropriate in backof Experiments on the LOB, Switchboard and WSJO corpora demonstrate that the technique greatly improves language i sparse training sets, and offers significantly improved complexity versus performance tradeoffs when compared with str

index Terms

Inspec

Controlled Indexing

computational linguistics linguistics word processing

Non-controlled Indexing

LOB Switchboard WSJO corpora ambiguous category membership backoff framework category bases models category based n grams language model perplexities part of speech classifications sequential re sparse training sets trigram models unseen word sequences word based language models

Author Keywords

Not Available

References

No references available on IEEE Xplore

Citing Documents

Using out-of-domain data to improve in-domain language models, Iyer, R.; Ostendorf, M.; Gish, H. Signal Processing Letters, IEEE On page(s): 221-223, Volume: 4, Issue: 8, Aug 1997 Abstract | Full Text: PDF (80)

indexed by 🗓 inspec Contact Us Privac

Copyright 2006 IE

[♦] View Search Results | ♦ Previous Article | Next Article >







Search

Full Text

Concept

Document ID

Recent Disclosures

Other

Prior Art Home

Support

Logout

Displaying records #1 through 6 out of 6

Result # 1 Relevance:

Vocabulary-Based Memory Compression

1985-05-01

IPCOM000063905D

The size or density of a document for memory storage and/or data transmission applica significantly reduced through replacing words in the document with look-up locations in table. Two vocabulary tables are employed with each accommodating 128 different ...

Relevance:

Relational KEY WORD Library

1989-07-01

IPCOM000035664D

En

Communications Log (COMLOG) or Systems Log (SYSLOG) electronic notebook type of used to store text data that describe activity against a processor test machine in an eng raised floor environment. Design engineers run test cases against test machines and ...

Result # 3

Relevance: (CC)

Procedures as a Representation for Data in a Computer Program for **Understanding Natural Language**

IPCOM000149027D

I 4. DESCRIPTIVE NOTES (Type O f report and Inclluaive datea) i , Ph.D. Thesis, Depart Mathematics. Ausust 1970 5. AUTHOR(§) (Last name, first name, initial) I Winograd, Te 1971 6. REPORT D A T E 7a. T O T A L NO. O F PAGES 7b. NO. O ...

Result # 4

Relevance:

IEEE Computer Volume 16 Number 6 -- NEW APPLICATIONS & RECEN RESEARCH

1983-06-01

IPCOM000131624D

NEW APPLICATIONS & RECENT RESEARCH * Dairy Co Op finds computers almost as im cows * Newspapers get help in automating their libraries * Computers go to Hollywood college to put a terminal in every dorm

Result # 5

Relevance:

L3 Lssd Latch Arrangement Schemes Which Permit 100% Cache Array Testing

1990-02-01

IPCOM000099598D

L3 latch arrangements have been used extensively to implement small CACHE memorie implementations resulted in the use of functional test patterns with little or no diagnost for testability as the use of LSSD patterns could not provide 100% coverage of ...

Result # 6

Relevance: 🛟

Genes and environment in language acquisition: a study of early voca syntactic development in twins

IPCOM000128114D

The goal of this thesis is to explore the contributions of heredity and environment to lar development using the twin method. This method consists of comparing monozygotic (I dizygotic (DZ) twins to see if MZs have a more similar course of development. Two ...

Displaying page 1 of 1 - << FIRST | < BACK | NEXT > ; LAST >>

Search query: word combination

New search | Modify this search | Search within current results





September 12, 2007

USPTO

Search	
--------	--

Full Text Concept Document ID Recent Disclosures

Other

Prior Art Home Support Logout

No records matched your search.

Perhaps you should try a less restrictive query.

Search query: (word combination) and corpus

New search | Modify this search

Conveight @ 2007 IP.com, Inc. All rights reserve

Web Images Video News Maps Gmail more v

Sign in

Google

+("word combination" or words or sentence or

Search

Advanced Search

Try uppercase "OR" to search for either of two town to details manage your web history

Web Results 1 - 10 of about 53 for +("word combination" or words or sentence or phrase or paragraph) +co

Tip: Save time by hitting the return key instead of clicking on "search"

[PDF] <u>LLC 18 3 249-259 fqq024 FINAL</u>

which the similarity of the meaning of words, sentences or paragraphs ... calculate the semantic similarity between two or more given segments. ... llc.oxfordjournals.org/cgi/reprint/18/3/249.pdf - Similar pages

[PDF] Towards automatic retrieval of idioms in French newspaper corpora File Format: PDF/Adobe Acrobat - View as HTML

sentence. For a large corpus segmented in many different documents, ... between words and text segments to screen the list of four-word segments from step ... cecl.fltr.ucl.ac.be/Downloads/Degand%20&%20Bestgen%202003.pdf - Similar pages

Topic identification and use thereof in information retrieval ...

Example segments include, but are not limited to, sentences, paragraphs, is defined as a paragraph, then the value of S2 for word-combination-i is the ... www.freepatentsonline.com/20030167252.html - 60k - Cached - Similar pages

[PDF] A Corpus-based Approach to the Chinese Word Segmentation

File Format: PDF/Adobe Acrobat

formation, word combination and the compound formation rule differ be- not merely a single single sentence or short paragraph. The disadvantage of ... edoc.ub.uni-muenchen.de/archive/00005662/01/Liu_Lezhong.pdf - Similar.pages

[PDF] Matrix: A statistical method and software tool for linguistic ...

File Format: PDF/Adobe Acrobat

sentences and paragraphs. We can distinguish the Matrix technique from corpus to calculate the expected frequencies of words, POS and semantic tags. ... www.comp.lancs.ac.uk/computing/users/paul/phd/phd2003.pdf - Similar pages

[PDF] Automatic discourse structure detection using shallow textual ...

determined by the contextual significance of the words or phrases which make it up. where P, Q, & R are the discourse segments in a sentence. ... linkinghub.elsevier.com/retrieve/pii/S1071581903002088 - Similar pages

[PDF] Extraction of Salient Textual Patterns: Synergy Between Lexical ...

recurrent word combination [8]. Collocation represents the re- coherence between the adjacent segments within the sentence. ...

ieeexplore.ieee.org/iel5/3468/28375/01268156.pdf - Similar pages

The Complete Bibliography of the ``Information Retrieval" journal ...

File Format: Unrecognized - View as HTML

co-occurrences of words in sentence-size or paragraph-size document identifier: the shortest word or word combination, that will locate the document. ...

nmis.isti.cnr.it/sebastiani/IRJ.bib - Similar pages

[PDF] Extracting Chinese Frequent Strings Without a Dictionary From a ...

File Format: PDF/Adobe Acrobat - View as HTML

+("word combination" or words or sentence or phrase or paragraph) +corpus +segments +... Page 2 of 2

(3) The probability of the word combination should be high. called breath groups or intonation phrases) in an input Chinese sentence. ... homepage.iis.sinica.edu.tw/JISE/2001/200109_07.pdf - Similar pages

[PDF] Navin Desai MSc Advanced Computer Science Supervisor: Mr. Yorick ... File Format: PDF/Adobe Acrobat - View as HTML technique to get around such spam filters, which make use of word combination methods. Such mails are generated by randomly combining sentences from various ... www.dcs.shef.ac.uk/intranet/teaching/projects/archive/msc2005/pdf/m4nd0.pdf - Similar.pages

1 <u>2</u> <u>3</u> <u>4</u> <u>5</u> <u>Next</u>

Download Google Pack: free essential software for your PC

+("word combination" or words or se Search

Search within results | Language Tools | Search Tips | Dissatisfied? Help us improve

©2007 Google - Google Home - Advertising Programs - Business Solutions - About Google

Web Images Video News Maps Gmail more •

Sign in

Google

+("word combination" or words or sentence or Search Professors

Try uppercase "OR" to search for either of two town wiedetails imanage your web history

Web Results 11 - 20 of about 53 for +("word combination" or words or sentence or phrase or paragraph) +c

Wordclass Tagging in BNC XML

Moreover, sometimes it is more appropriate to tag a word combination as These are the segments usually demarcated by (1) spaces and (2) sentence ... www.natcorp.ox.ac.uk/XMLedition/URG/posguide.html - 214k - Cached - Similar pages

Search engine methods and systems for generating relevant search ... Example segments include, but are not limited to, sentences, paragraphs, [0055] One method to compute the expected usage of significant words in data ... www.freshpatents.com/Search-engine-methods-and-systems-for-generating-relevantsearch-results-and-adverti... - 100k - Cached - Similar pages

[PDF] A Corpus-based Approach to the Chinese Word Segmentation Chinese Penn Tree Bank 225235 words and phrases tagged for POS/function 2003]. which calculate the probability for the whole sentence. We ... deposit.ddb.de/cgi-bin/doksery?idn=980925770& dok_var=d1&dok_ext=pdf&filename=980925770.pdf - Similar pages

Топик: Modern English Word-Formation - BestReferat.ru - Банк ...- [Translate this page] Similarly, he can conceive of sentences containing the word GRACIOUSLY'S In that way, "a legal term is a word or a word combination which stands for ... www.bestreferat.ru/referat-39172.html - 144k - Cached - Similar pages

[PDF] Broadcast news processing: Structural Classification ...

File Format: PDF/Adobe Acrobat - View as HTML across the entire corpus are taken to calculate the density of each named paragraph of 50 words or a sentence of 15 words, it must adhere to these ... www.dcs.shef,ac.uk/~balakkvj/final_finalThesis_single_spaced.pdf - Similar pages

Synchronized pattern recognition source data processed by manual ... or other image-based pattern recognition, such as computer aided diagnosis for Segmentation options may include paragraph, sentence, phrase, or word. ... www.freepatentsonline.com/20060149558.html - 402k - Cached - Similar pages

[PDF] Controlled Languages in Software User Documentation Dina ... File Format: PDF/Adobe Acrobat - View as HTML di erent levels: word, phrase, clause, sentence, paragraph, and text. During captures the adequacy aspect, while the word combination search accounts ... www.diva-portal.org/diva/getDocument?urn_nbn_se_liu_diva-4637-1__fulltext.pdf -Similar pages

[PDF] Introduction to Natural Language Processing / Computational ...

File Format: PDF/Adobe Acrobat

trees from this corpus, find instances of verb phrase conjunctions With the help of your word and sentence tokenizers, compute the ARI scores ... www.ldc.upenn.edu/sb/nltk-book-part1.pdf - Similar pages

[PDF] Proceedings of the...

File Format: PDF/Adobe Acrobat

+("word combination" or words or sentence or phrase or paragraph) +corpus +segments +... Page 2 of 2

ally taken to be any word combination (adjacent or. otherwise) that has some feature (syntactic, each **phrase** they **calculate** the probability of seeing ... acl.ldc.upenn.edu/W/W07/W07-11.pdf - <u>Similar pages</u>

[poc] Integrating ESL and Translation into Second Language: the Marriage ... File Format: Microsoft Word - View as HTML

Table 4-7 shows the category of each kind of word combination in (12): We calculate the similarity measure for each sentence in TM DB (2) to find out ...

www.swan.ac.uk/cals/staff/shei/phd2/thesis_revised2.doc - Similar pages

Previous 1 2 3 4 5 Next

+("word combination" or words or se Search

Search within results | Language Tools | Search Tips

©2007 Google - Google Home - Advertising Programs - Business Solutions - About Google

Web Images Video News Maps Gmail more

Sign in

Google

+("word combination" or words or sentence or

Search

Advanced Search

Try uppercase "OR" to search for either of two wo wide tails manage your web history

Web Results 21 - 30 of about 53 for +("word combination" or words or sentence or phrase or paragraph) +c

[PDF] Multimedia Human-Computer Interface

File Format: PDF/Adobe Acrobat

D(i) is any DU (e.g., a phrase,. sentence, or "paragraph"). mands (a word or a word combination) is predefined for this. purpose. ...

www.hindawi.com/journals/ASP/si/0012004009.pdf - Similar pages

[PDF] COMPARATIVE EVALUATION OF MODULAR AUTOMATIC SUMMARISATION SYSTEMS ...

File Format: PDF/Adobe Acrobat

BNC - British National Corpus CAST - Computer-Aided Summarisation Tool Quite often these words appear in indicating phrases, and sentences ... home.wlv.ac.uk/~in6093/papers/PhD/Thesis.pdf - Similar pages

[PDF] Computational Linguistics in the Netherlands 2004

File Format: PDF/Adobe Acrobat

When using bad segments as the background corpus, one would form part of a recurrent word-combination in one way or another." (p. 102). Bybee ... lotos.library.uu.nl/publish/articles/000146/bookpart.pdf - Similar pages

The Project Gutenberg Etext of The 1913 Webster Unabridged ...

A word combination containing the smaller by headword (or a morphological ... These segments may be separate, as in the separate paragraphs starting ... www.gutenberg.org/dirs/etext96/pgw050ab.txt - Similar pages

[PDF] Hanzi, Concept and Computation: A Preliminary Survey of Chinese ... where along a continuum between Free-word Combination and Idioms. such as words, phrases and sentences, and allows for structure sharing via co- ... deposit.ddb.de/cgi-bin/dokserv?idn=979621003& dok_var=d1&dok_ext=pdf&filename=979621003.pdf - Similar pages

[PDF] arXiv:cmp-lg/9609008 v1 25 Sep 1996

File Format: PDF/Adobe Acrobat

random sentence from a corpus. Suppose it contains n words, w Not only is this example commonly used as a fixed word combination, but its meaning ... arxiv.org/pdf/cmp-lg/9609008 - Similar pages

[PDF] 國立清華大學碩士論文

File Format: PDF/Adobe Acrobat - View as HTML

It can show example sentences of a key. word/phrase when one types in a word or phrase for searching. The corpus in. TOTAL recall is composed of articles ... www.fl.nthu.edu.tw/thesis/tesol/915260.pdf - Similar pages

[PS] Designing Statistical Language Learners: Experiments on Noun ...

File Format: Adobe PostScript

random sentence from a corpus. Suppose it contains n words, w Not only is this example commonly used as a xed word combination, but its meaning would ... www.ics.mq.edu.au/~mark/mclthesis.ps.gz - Similar pages

+("word combination" or words or sentence or phrase or paragraph) +corpus +segments +... Page 2 of 2

(PS) Untitled

File Format: Adobe PostScript

all sentences of up to twenty words of the corpus (in as far as these were annotated To compute the phrase, entropy score of a tuple (P,N) (e.g. (in ... odur.let.rug.nl/~vannoord/alp/midterm.ps - Similar pages

[PDF] Term-driven E-Commerce

File Format: PDF/Adobe Acrobat

words, phrases or sentences. In TE-Commerce, brand names and product types could right segments with which they appeared:. 268 computer. 259 medien ... edoc.ub.uni-muenchen.de/archive/00006367/01/Rolletschek_Gerhard.pdf - Similar pages

<u>Previous 1 2 3 4 5</u> **Next**

+("word combination" or words or se Search

Search within results | Language Tools | Search Tips

©2007 Google - Google Home - Advertising Programs - Business Solutions - About Google

Web Images Video News Maps Gmail more •

Sign in

Google

+("word combination" or words or sentence or Search Refresses

Try uppercase "OR" to search for either of two town to details imanage your web history

Web Results 31 - 39 of 39 for +("word combination" or words or sentence or phrase or paragraph) +corpus

IPDEJ UNIVERSIDAD DE EXTREMADURA FACULTAD DE POLITÉCNICA DPTO. DE LENGUA ...

File Formal: PDF/Adobe Acrobat

context of words in sentences and phrases, genres could not be identified formulating phrases and paragraphs based on concept or event development. ... dialnet.unirioja.es/servlet/fichero_tesis?articulo=1349070&orden=0 - Similar pages

[PDF] Proceedings of the Linguistic Annotation Workshop

File Format: PDF/Adobe Acrobat

of Computer Science. A. Belz and S. Varges, editors: 2005. Corpus Linguistics 2005 </sentence>. Dependency of word combination is annotated ... acl.ldc.upenn.edu/W/W07/W07-15.pdf - Similar pages

[PDF] Section 4. Intelligent Technologies in Control, Design and ...

File Format: PDF/Adobe Acrobat

Syntactic relations between words are ignored, and words from different sentences or even paragraphs are. analyzed. In the paper [3] malapropism detection ... www.foibg.com/conf//TA2005/KDS05-Book-Sec4.pdf - Similar pages

[PDF] XIII-th International Conference Knowledge-Dialogue-Solution ...

File Format: PDF/Adobe Acrobat

In turn, each unit for a particular word combination accumulates necessary A sentence or even a text paragraph is taken as an input for searching. The ... www.foibg.com/conf/proceedings/KDS07-Book2.pdf - Similar pages [More results from www.fcibg.com]

[PDF] Variations on Language Modeling for Information Retrieval

File Format: PDF/Adobe Acrobat

in a corpus and especially for phrases. An alternative for the tests ponents (stems) on the basis of word combination rules for Dutch and a lexicon. ... www.ctit.utwente.nl/library/phd/kraaij.pdf - Similar pages

**This is a COPYRIGHTED Project Gutenberg Etext, Details Below ...

In the event that you determine to bring a writ of habeas corpus proceeding some low fellow (I'd ordinarily use a four-word combination we use and ... www.gutenberg.org/dirs/etext01/efpap10.txt - 454k - Cached - Similar pages

Padilla Convicted On All Counts, But Was A Trial Necessary?

... cover dvd mcbeal dictionary words with sentences bourbon street shrimp guest name net online order site word combination generator h2200 cradle ... 36c.8sila . cn/ - Similar pages

preschool theme days convert text to pdf linux york hotel in san ...

... andaluz mijas sumer set computer overheating solutions karol adamiecki tx 911 ignition switch word combination dictionary grinder accessories keak ... smnw.wzgcf . cn/kufpmek.html - Similar pages

ACTION UKRAINE REPORT - AUR: November 2005

+("word combination" or words or sentence or phrase or paragraph) +corpus +segments +... Page 2 of 2

The word combination "free, honest and Maybe you have an old computer that can be fixed and installed in the classroom. Internet providers can donate ... action-ukraine-report.blogspot.com/2005_11_28_action-ukraine-report_archive.html -Similar pages

In order to show you the most relevant results, we have omitted some entries very similar to the 39 already displayed.

If you like, you can repeat the search with the omitted results included.

Previous 1 2 3 4

+("word combination" or words or se

Search within results | Language Tools | Search Tips

©2007 Google - Google Home - Advertising Programs - Business Solutions - About Google